A STEP-SIZE SELECTION PROCEDURE FOR EQUALITY CONSTRAINED OPTIMIZATION *

Jean Charles GILBERT

International Institute for Applied Systems Analysis
A - 2361 Laxenburg (Austria)

Résumé. Dans cet article, on propose une méthode généralisant à l'optimisation avec contraintes d'égalité, une technique de sélection de pas, dite de Wolfe, qui a fait ses preuves en optimisation sans contrainte. Pour les algorithmes du type quasi-Newtonien, celle-ci semble assez naturelle puisqu'elle permet d'assurer facilement la définie positivité des métriques locales à chaque itération et donc le caractère descendant des directions de recherche. Toutefois, on sait que cette technique ne peut pas être étendue aux méthodes quasi-Newtoniennes en optimisation avec contraintes. On montre ici qu'une généralisation est possible dans le cadre des méthodes sécantes réduites.

Abstract. This paper proposes a generalization to equality constrained optimization of Wolfe's step-size selection procedure, which is used with success in unconstrained optimization. This one appears rather natural for quasi-Newton methods because it allows to maintain easily the positive definiteness of the matrices correcting the steepest descent direction and therefore assures the descent property to search directions. However, this technique is known not to be usable for quasi-Newton methods in constrained optimization. We show here that a generalization can be made in the framework of reduced secant methods.

* Work supported by the FNRS (Fonds National de la Recherche Scientifique) of Belgium.

1. Introduction

In this report, we deal with the following equality constrained optimization problem:

$$\min \{ f(x) : x \in \omega \, , \, c(x) = 0 \} \, , \tag{1.1}$$

where ω is a convex open set in $\mathbb{R}^n$ and the functions $f : \omega \longrightarrow \mathbb{R}$ and $c : \omega \longrightarrow \mathbb{R}^m$, $m < n$, are supposed smooth. We shall suppose that the $m \times n$ Jacobian matrix $A(x) := \nabla c(x)$ has full rank m for all x in ω and that ω contains a local solution x_* of problem (1.1), which with its associated Lagrange multiplier λ_* satisfies the standard sufficient conditions of optimality (see Fletcher (1981)):

$$\begin{cases} c(x_*) = 0 \, , \\ \nabla f(x_*) + A(x_*)^T \lambda_* = 0 \end{cases} \tag{1.2}$$

and

$$G_* := Z(x_*)^{-T} \, L_* \, Z(x_*)^- \ \text{ is positive definite} \, . \tag{1.3}$$

In (1.3), L_* is the Hessian according to x of the Lagrangian $l(x,\lambda) := f(x) + c(x)^T\lambda$ at (x_*,λ_*) and $Z(x_*)^-$ is a basis of $N(A(x_*))$, the kernel of $A(x_*)$, i.e. an $n \times (n-m)$ matrix whose columns form a basis of $N(A(x_*))$. We shall suppose that such a basis exists at each point x in ω in such a way that its dependence on x is smooth. We have

$$A(x) \, Z(x)^- = 0 \ \text{ in } \mathbb{R}^{m\times(n-m)} \text{ for all } x \text{ in } \omega \, . \tag{1.4}$$

We shall also need to do displacements in the complementary space $R(A(x)^-)$ of $N(A(x))$. We shall also suppose that the right inverse $A(x)^-$ of $A(x)$ is a smooth function of x. We have

$$A(x) \, A(x)^- = I \ \text{ in } \mathbb{R}^{m\times m} \text{ for all } x \text{ in } \omega \, . \tag{1.5}$$

This formalism (with the matrix $Z(x)$ introduced later) is due to Gabay (1982).

We shall focus in this paper on the following class of reduced secant methods to compute iteratively a solution of problem (1.1) (see Coleman and Conn (1982)): starting from a point x_k, the next iterate x_{k+1} is obtained by

$$y_k := x_k - A(x_k)^- \, c(x_k) =: x_k + r_k \, , \tag{1.6}$$

$$x_{k+1} := y_k - Z(y_k)^- \, H_k \, g(y_k) =: y_k + t_k \, . \tag{1.7}$$